

Original communication

A grey zone approach for evaluation of 15 short tandem repeat loci in sibship analysis: A pilot study in Indian subjects

Rajiv I. Giroti M.Sc.^{a,*}, Sunita Verma M.Sc.^a, Kulwant Singh M.Sc.^a, Rohit Malik M.Sc.^a,
Indu Talwar Ph.D. (Professor and Chair)^b

^a Central Forensic Science Laboratory, Plot No. 2, Dakshin Marg, Sector 36-A, Chandigarh 160 036, India

^b Department of Anthropology, Panjab University, Chandigarh 160 014, India

Received 8 May 2006; received in revised form 8 August 2006; accepted 10 August 2006

Abstract

The evaluation of 15 STR loci Applied Biosystems Identifier kit for sibship determination in Indian subjects is reported. Cumulative sibship indices (CSIs) calculated following standard methods in sibling pairs and non-sibling pairs, showed mean values comparable to other reports. Mean CSI value in sibling group was higher than in corresponding non-sibling group. Moderately high value of CSI in one of the non-sibling pairs and a very low likelihood ratio favoring non-relatedness in a known sibling pair did not allow binary decision about sibship status. To deal with this problem a grey zone approach has been applied to sibship test. It is concluded that the 15 loci STR kit can be reliably used for inferring sibship between pairs of individuals by defining a grey zone of a sibship test as an area of likelihood ratio values where the discriminatory performance is insufficient.

© 2006 Elsevier Ltd and AFP. All rights reserved.

Keywords: Forensic science; Sibship; STR loci; Grey Zone; Human identification; Cumulative sibling index

1. Introduction

Forensic DNA analysis is traditionally a science of comparison. In cases of personal identification of unknown human remains, a direct comparison can be carried out between the DNA profiles developed from the remains with those obtained from the personal articles of the deceased.¹ While direct comparison provides the most meaningful conclusion, a practical problem with the use of personal articles as references is that these articles may be unavailable for study or they may not yield sufficient DNA for analysis. Hence the Forensic DNA analysts mainly rely on reference samples of the close biological relatives of the deceased for indirect comparison. Parents of the deceased in the absence of mutation, offer a rich source of genetic information for inferring correct relationship.

In instances where parentage analysis is not feasible the investigators turn to the next closest genetic kinship, i.e., of the full siblings, where allele sharing by descent is an observed event. Sibship indices are used in determining a hypothesized relationship between two persons.² This approach uses a likelihood ratio to evaluate the support of the evidence under the hypothesis that the evidence profile is from the sibling or from an unrelated person.

STR multiplexes assays are now the dominant forensic human identification technology.³ A core set of 15 STR markers are now being used worldwide and due to high discrimination power are supposed to maximize the relative probabilities of allele sharing by descent versus allele sharing by chance but their accuracy in discriminating siblings from non-siblings is unclear. Previous studies using various STR locus combinations reflect that it is difficult to determine exactly how many loci are necessary to achieve a certain level of power in sibship inferences and highly impossible to provide a universal CSI cut-off point for

* Corresponding author. Tel.: +91 172 2621441; fax: +91 172 2605923.
E-mail address: righiroti@yahoo.com (R.I. Giroti).

demarcating siblings from unrelated group.^{4,5} These studies also suggest for creation of an initial reference standard in every DNA laboratory on which the decision to apply the test convincingly can be taken. The present case-control sibship study in Indian subjects was designed with an aim of generating suitable values of CSI using the 15 panel STR multiplex systems. We have examined the role of CSI in predicting sibship in Indian subjects across the guidelines established for degree of certainty based on the calculated LR values.⁶ We determined conventional sensitivity and specificity, as well as positive and negative likelihood ratios (LR+ and LR–). This has allowed us for defining the grey zone of sibship test for screening discrimination.

2. Materials and methods

In a pilot study, we selected 33 unrelated volunteer family units, consisting of a mother, father and two siblings. 33 pairs of unrelated non-siblings were taken for control studies. The whole blood samples were archived on FTA cards (Whatman Biosciences® UK).

2.1. DNA purification

The DNA purification steps were performed as per the manufacturer's instructions (Whatman Biosciences® UK).

2.2. DNA amplification

AmpFISTR identifier™ multiplex assay (Applied Biosystems, Foster City, California) for 15 autosomal, co-dominant and unlinked STR loci was used for the present study. The amplification reactions were performed according to the manufacturer's instructions. The samples were amplified in a 2400 thermal cycler (PE Applied Biosystems, Foster City, California).

2.3. Separation and detection of amplified products

Amplified product (1.5 µl) was dispensed into a tube containing 24 µl of Hidi Formamide (Applied Biosystems, Warrington, UK) and 0.5 µl of Liz-500 size standard (Applied Biosystems, Warrington, UK). The tube was heated to 95 °C for 3 min and then snap chilled in a portable chiller for 3 min. Electrophoresis was carried out on an ABI 310 GeneticAnalyzer (Applied Biosystems, Foster City, California) with GeneScan 2.1 software. Genotypes were determined by comparing the size of the unknown fragment to the allelic ladder that was run in parallel.

2.4. Statistics and evaluation

Once the genotypes for the 33 family units were determined, we conducted paternity tests to establish that the two children of each family unit were biological children of both parents. Sibship index values were then investi-

gated in 33 true sibling pairs and compared with those of other 33 non-related random pairs.

Paternity and sibship indices were calculated using 'PATCAN' software.⁷ Average power of discrimination values fall within the range of 0.718 and 0.941 for 15 STR loci in our in-house population data. The typical paternity index is 549313, while the minimum cut-point established for determining paternity is 1000. Although no consensus has been reached on CSI cut-off value to use with sibship test, the cut-off level of 1 was used to define a positive test result. Sensitivity and specificity were determined at this level. Test sensitivity was calculated as true positive tests per total sibling pairs tested. Test specificity was calculated as true negative tests per non-sibling pairs tested. Both these parameters were expressed as percentages. Likelihood ratio for positive test result (LR+) and Likelihood ratio for negative test result (LR–) were defined in terms of determined sensitivity and specificity: $LR+ = \text{sensitivity} / (1 - \text{specificity})$ and $LR- = 1 - \text{sensitivity} / \text{specificity}$. Two cut-off points one associated with the minimal desirable value of LR(+) and the other maximal desirable value of LR(–) were identified delimiting the grey zone (area of inconclusive CSI values) as described by Joël Coste and Jacques Pouchot.⁸

3. Results

The results of cumulative sibship indices (CSIs) for both sibling and random pairs are shown in Fig. 1. The lowest CSI value of 0.0000003 was noted in the non-sibling pair, while the highest value of 116416023 was illustrated in a sibling pair. For sibling pairs both mean and median CSI values were significantly higher when compared to non-sibling pairs. In the sibling group 31 cases (93.94%) had CSI values greater than 1. Two known sibling pairs (6.06%) illustrated CSI < 1 with one of them ranked as low as 0.0019. Thirty cases (90.9%) of the random pairs had CSI values less than one. Fortunately one non-sibling pair had CSI as high as 17.2. Fig. 2 shows the sensitivity and the specificity of the test at CSI cut-off level of 1. The likelihood ratio indicating the value of the test for increasing certainty about a positive judgment (LR+) is 10.3 (CSI cut-off point = 1). The likelihood ratio indicating the value of the test for increasing certainty about a negative opinion (LR–) is 0.067 (CSI cut-off point = 1). A grey zone keeping under surveillance the inconclusive CSI values for sibship test is shown in Fig. 3.

The distribution of allele sharing for the sibling and the non-siblings is shown in Fig. 4. There were a total of 495 observations (15 loci × 33 pairs) for each group. The sibling pair with lowest CSI value (0.0019) conspicuously did not show the phenomenon of '2- allele sharing' at any of the loci. Zero allele sharing was observed at three loci in this pair. There were nine non-sibling pairs that showed no 2 alleles sharing at any of the loci. There was a perceptible '0-allele sharing' at 12 loci in one of the non-sibling pairs in the midst of the lowest CSI.

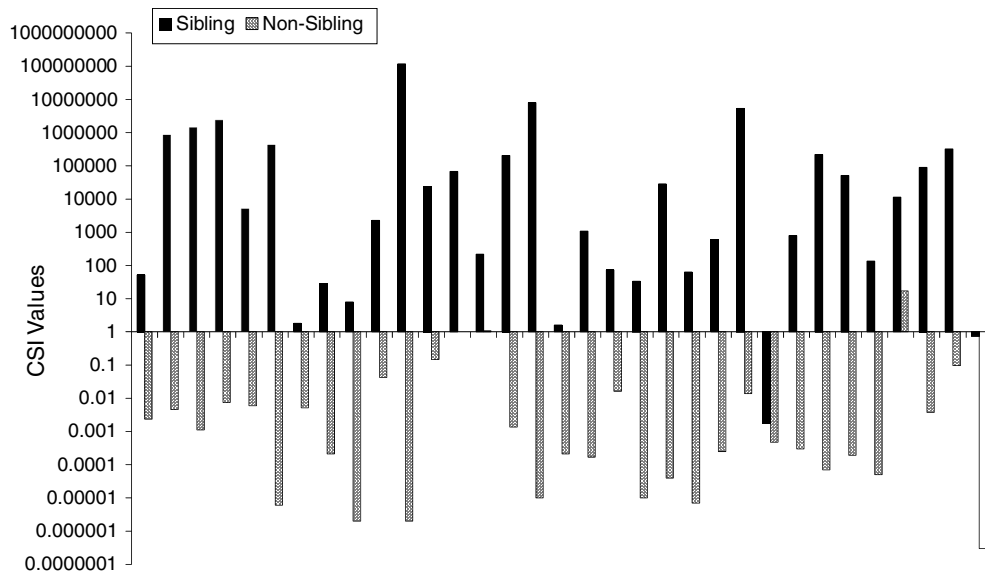


Fig. 1. Cumulative sibship indices (CSIs) for sibling and non-sibling pairs.

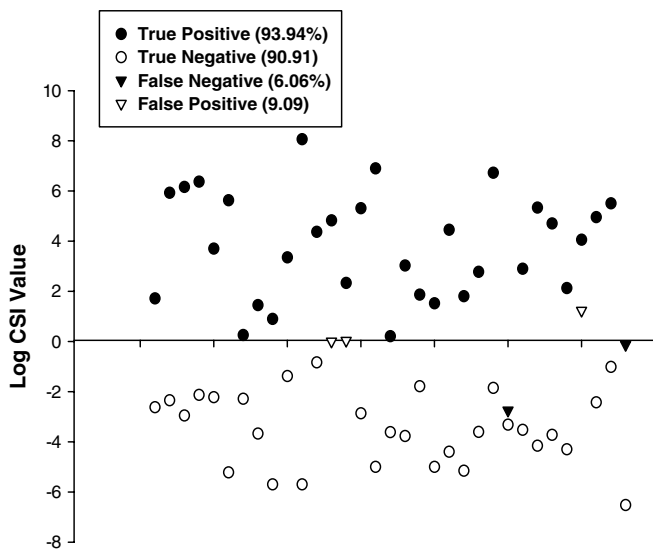


Fig. 2. Sensitivity and specificity of sibship testing at threshold value of CSI = 1.

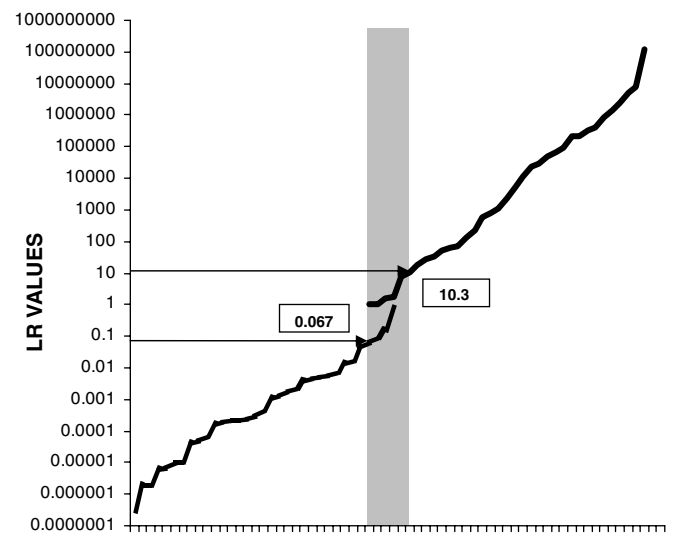


Fig. 3. Construction of grey zone for sibship test for different values of CSI, (LR+) and (LR-) associated with CSI cut-off point = 1 have been plotted delimiting the area.

4. Discussion

The usefulness of DNA markers for kinship determinations has received much attention in forensic literature. Useful statistical formulas have been established for calculation of sibship index based on the population frequencies of the shared alleles at multiple DNA loci to indicate the degree of relatedness in any two persons.² With every new DNA marker system making its way in forensic laboratories different authors have prescribed threshold CSI for inferring the likelihood of probable sibship. Gaytmenn et al.⁵ have recently reported significant risks of false inclusion associated with this approach when 3, 6 and 9 loci

STR systems are used. Tzeng et al.⁴ using 15 loci STR panel established a presumptive cutting point of CSI at ≥ 3.0 for sibship determination. For Reid et al.⁹ while testing ABI Identifiler multiplex system (15 STR loci) the likelihood ratio of probable relationship for the non-sibling pairs did not increase above the prescribed threshold value of 1. Thus no consensus is possible on what CSI cut-off point should be used to discriminate siblings from non-related individuals. In the present study CSI as low as 0.0019 in a known sibling pair and CSI as high as 17.2 in one of the random pairs presented a complicated episode for defining an optimal cut-point for perfectly discriminating between subjects with and without sibship. With CSI

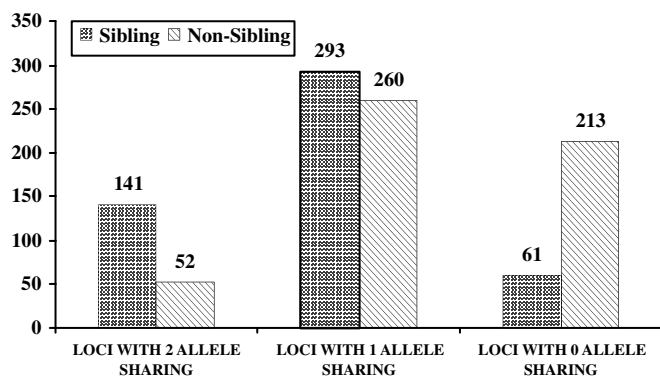


Fig. 4. Distribution of allele sharing for the sibling and the non-sibling group.

value set at 1 for defining a positive sibship, the test has correctly identified 31 out of 33 siblings (Sensitivity 93.94%). The number of false negatives is low (6.06%). Having a high sensitivity is not necessarily a good thing since a sensitivity of 100% can always be achieved by using the simple strategy of always producing a positive result. Specificity, on the other hand in our study is concerned with how good the test is at correctly identifying people who are non-siblings. While settling for a threshold value just above 17 will raise the specificity of the test to 100% the test tends to lose its sensitivity generating about 15% false negatives. From forensic perspective, it is very important not to miss a diagnosis and have false negatives while examining a criminal case hence a threshold value for maintaining a high sensitivity may be adopted. It is equally important not to create false positives as in victim identification cases and the DNA analyst may give greater emphasis to assessing the test with respect to its specificity and seek a threshold value for higher specificity. Therefore assessing a new kit or test will not be simple under the given constraint of single CSI cut-off value. We have therefore constructed a three-zone partition for sibship test to avoid the constraint of a 'black or white' decision as proposed by Joël Coste and Jacques Pouchot.⁸ This has allowed a partitioning of our CSI results into definitely negative area, most certainly positive area and a grey area (requiring another genetic test). Given these areas it can be concluded that most accurate estimate of CSI ranges from 0.067 to 0.0000003 for negative findings, from 10.3 to 116416023 for positive findings and an inconclusive area spanning between 0.067 and 10.3. Although the problematic CSI values of 0.0019 (in the known sibling pair) and 17.2 (in unrelated pair) are still located outside the grey area, the stratified values of CSI are more useful than single cut-off values reported by other authors.

It is postulated that sharing of two alleles at multiple loci is an informative estimator while accounting for sibship. Full siblings are expected to share zero or two alleles identical by descent at a given locus with an equal probability of 0.25 and share a single allele with a probability of 0.5. In this study impressive sharing of two alleles at

numerous loci was noted for the sibling group. The percentage of zero, one and two allele sharing was 12.32%, 59.19% and 28.48% in this group. High genetic relatedness was encountered in one of the sibling pairs where sharing of two alleles was observed at 10 out of 15 loci. The high level of relatedness among this sibling pair could be rooted in the fact that parent genotypes together featured nine loci in a homozygous state and extensive one allele match at 6 loci. A fortuitous situation involving a sibling pair also arose where the phenomenon of two allele sharing was absolutely lacking. The very low CSI in this sibling pair can be attributed to lack of two allele sharing phenomenon at all of the 15 genetic loci tested. There was also a circumstance in this study where one of the unrelated pairs featured two allele sharing phenomenon at 5 loci consequently resulting in highest CSI (17.2) among the unrelated pairs. Investigation of individual loci in non-related pair with high CSI showed that this sudden increase in CSI could not be associated with single chance sharing of very rare allele. Despite of these anomalies the distribution of loci number for two-allele sharing for random and sibling group showed a precise polarization. A similar polarization was observed for sharing of 0 alleles. However no such polarization was noted for single allele sharing where the percentage sharing for siblings and non-siblings was 59.19% and 52.52%, respectively. Tzeng et al.⁴ and Reid et al.⁹ have also reported this phenomenon. The extreme polarization observed in the degree of zero and 2 allele (s) sharing across various loci in sibling and non-sibling groups indicates that these loci maximize the relative probability of allele sharing by descent versus allele sharing by chance.

The present study is based on samples from Hindu families. The caste system reflects occupational and socio-religious hierarchies within this community.^{10,11} While caste endogamy is universal in this religion a degree of non-uniformity in attitude to close kin marriages exists. The Aryan Hindus of the northern India prohibit marriage between biological kin and there is culturally defined lineage exogamy.¹² In comparison Dravidian Hindus of south India strongly favor marriage among first cousin.^{13,14} These social variables have not been taken into account in the study. With strong evidence of genetic stratification¹⁵ and random or preferential inbreeding it will be interesting to assess the validity of these markers in kinship studies under the influence of various social variables. The present study represents initial findings of a small sample size completely without stratification. Variation in the populations of the subjects may have introduced bias. Nevertheless it is interesting to note that on the basis of CSI information provided by the 15 STR loci one can approach the problem of assessing the sibship existing among samples of unknown relatedness. The construction of grey zone will lead to more reliable identification and will help to satisfy the increased judicial regulations that require experts to define error rates for the tests.

References

1. Sasaki M, Shiono H, Fukushima T, Shimizu K. Human Identification by genotyping of personal articles. *Forensic Sci Int* 1997;**90**: 65–75.
2. Wenk RE, Traver M, Chiafari FA. Determination of sibship in any two persons. *Transfusion* 1996;**36**:259–62.
3. Kimpton CP, Oldroyd NJ, Watson SK, Frazier RR, Johnson PE, Millican ES, et al. Validation of highly discriminating multiplex short tandem repeat amplification systems for individual identification. *Electrophoresis* 1996;**17**(8):1283–93.
4. Tzeng CH, Lyou JY, Chen YR, Hu HY, Lin JS, Wang SY, et al. Determination of sibship by PCR-amplified short tandem repeat analysis in Taiwan. *Transfusion* 2000;**40**:840–2.
5. Gaytmenn R, Hildebrand DP, Sweet D, Pretty IA. Determination of the sensitivity and specificity of sibship calculation using AmpFISTR Profiler Plus. *Int J Legal Med* 2002;**116**:161–4.
6. Evett IW. *Strength of evidence in likelihood ratios. The use of statistics in forensic science*. New York: Ellis Horwood; 1991.
7. Riancho JA, Zarrabeitia A. window based software for common paternity and sibling analysis. *Forensic Sci Int* 2003;**135**(3):232–4.
8. Coste Joël, Pouchot Jacques. A grey zone for quantitative diagnostic and screening tests. *Int J Epidemiol* 2003;**32**:304–13.
9. Reid TM, Wolf CA, Kraemer CM, Lee SC, Baird ML, Lee RF. Specificity of sibship determination using the ABI Identifiler Multiplex System. *J Forensic Sci* 2004;**49**(6):1262–4.
10. Bhasin MK, Walter H, Danker-Hopfe H. The Distribution of Genetical, Morphological and Behavioural Traits among the Peoples of Indian Region. Delhi: Kamla-Raj; 1992, p. 14–35.
11. Gadgil M, Joshi NV, Manoharan S, Patil S, Prasad UVS. Peopling of India. In: Balasubramanian D, Appaji Rao N, editors. *The Indian Human Heritage*. Hyderabad Universities Press; 1998. p. 100–29.
12. Kapadia KM. *Marriage and family in India*. 2nd ed. Calcutta: Oxford University Press; 1958. pp. 117–137.
13. Appaji Rao N, Savithri HS, Bittles AH. A genetic perspective on the South Indian tradition of consanguineous marriage. In: Driesen CV, Nandan S, editors. *Austral-Asian encounters*. New Delhi: Prestige Books; 2002. p. 326–41.
14. Rao PSS. Inbreeding in India-concepts and consequences. In: Lukacs JR, editor. *The People of India*. New York: Plenum; 1984. p. 239–68.
15. Bamshad MJ. Genetic evidence on the origins of Indian caste populations. *Genome Res* 2001;**11**:1574–85.